Reference Sources Review

Jonathan Leff

April 17, 2010

Archives and Manuscripts

LIBR256-01 – Spring 2010

School of Library and Information Science – San Jose State University
Abstract

A study of five different websites designed to allow users to search archival collections reveals a variety of ways of providing access to archival materials. Web-based tools provide information via simple and advanced search capabilities and more in-depth systems such as MARC21 encoding. Some sites have the capability of providing online digital reproductions of materials in archival collections. Being able to provide easy to use tools and ready access to a large amount of archival materials is important in that it allows researchers to focus on their work and find the materials that will support their work in the easiest and most efficient manner possible.
Introduction

We live in an age when more people have access to the holdings of archives and special collections than ever before due to tools such as online catalogs and databases. Some of these tools are more robust in scope, while others facilitate ease of use and searching through the design of their interfaces. In this paper I explore five online tools for providing access to information about archival collections, including some that have access to digital resources. These sites are WorldCat: UC Berkeley Libraries, the National Register of Archives (a catalog of holdings in the British National Archives), the Online Archive of California, the National Union Catalog of Manuscript Collections (a project of the Library of Congress), and the Repositories of Primary Sources, which is a stand-alone site affiliated with the University of Idaho Library. In this paper I describe the sites, conduct searches, and evaluate the sites on ease of use and efficiency in finding archival resources online.

Scope And Content

WorldCat: UC Berkeley Libraries. (2010) is a pilot program in which UC Berkeley’s online catalog is linked via WorldCat to online catalogs throughout the UC system, and also worldwide. WorldCat is a union catalog that is maintained by the Online Computer Library Center (OCLC). The current (non-pilot) version is Melvyl, which is an online union catalog that is limited to the UC library system. According to its home page (OCLC, 2010b), WorldCat is connected to over 10,000 online library catalogs. By participating in WorldCat, the UC Berkeley Libraries have in effect joined their catalog with all of the libraries registered with WorldCat, meaning that UC Berkeley’s catalog is searchable from any other library that participates in the WorldCat system.

The National Register of Archives (NRA, 2010) is the British National Archives’ web portal and archives catalog, which provides information on records maintained by the National Archives (2010), as well as information as to the repositories where particular records are stored throughout Great Britain. The National Archives is comprehensive in that it claims to contain 1,000 years of UK government records as well as 270,000,000 accessible online documents. The NRA has five indexes to the records: the Business Index, the Organisations Index, the personal index, the Families and Estates Index and The Diaries and Papers Index. The front page of the National Register of Archives contains four separate search portals, some of which allow the user to search up to two indexes simultaneously.

The Online Archive of California (OAC, 2009) is run by the California Digital Library, which in turn operates under the auspices of the Regents of the University of California. It operates in a similar manner to the National Register of Archives, in that it is web portal that provides information about the holdings in the institutions participating in the program. These institutions include museums, public libraries, libraries affiliated with public and private universities, and collections associated with research institutions such as NASA Ames Research Center and public institutions such as California State Parks. The OAC provides information not only on the locations of physical collections, but also provides access to online collections where available.
The National Union Catalog of Manuscript Collections (NUCMC, 2009) was created by the Library of Congress to coordinate cataloging of and access to cataloging collections throughout the United States and Canada. It was created in 1959 as a print catalog and has been entirely online since 1993. Repositories send descriptions of their collections to the Library of Congress which then creates records in WorldCat using MARC21 encoding.

The Repositories of Primary Sources (Abraham, 2009) was compiled by Terry Abraham and is linked to the Special Collections section of the University of Idaho Library under the heading “Other Sources”. It acts as a web portal to library and archive websites around the world. The links are divided in multiple ways, with the primary ways being according to region, alphabetically according to country and province and with a link titled “Additional Lists” which contains links to other resources that are not included in the geographically sorted links. The site also features an “Integrated Index/List” which contains an alphabetical list of all the links in a scrolling frame on the left-hand side of the page.

Evaluation

I tested the WorldCat: UC Berkeley Libraries system by doing a simple search for “Halldor Laxness”, who is an Icelandic writer. There is a drop-down menu to the right of the search field which allows the user to focus results based on their availability at UC Berkeley only, the UC system as a whole, and libraries worldwide. I chose the third option. The link for “Archival Materials” was not visible under the heading for “Format” in the sidebar, though I found it by clicking “search more”. The one result returned was titled “Series I., General Correspondence, 1922-1977 Records.”, and contained a long list of people’s names, including “Halldor Laxness”. I guessed correctly that the title of the series was a hyperlink, and when I clicked on it, I was brought to a page that contained a summary indicating that this series was the central editorial files of Alfred A. Knopf, Inc. which included author files and correspondence. Beneath the summary was a section that stated that the materials were not available in the UC Berkeley Library, but which allowed the user to search all of the WorldCat participating libraries to find out where it was located. I did so and found out that it is held at the University of Texas at Austin, Harry Ransom Center. Clicking on the name of the University of Texas brought me to that library’s WorldCat page which contained a more detailed description of the archive including call number, size of collection in boxes and square feet, and also the fact that it has an unpublished finding aid.

I attempted my first search of the NRA by searching under the “Corporate Names” link, which combines the Business Index and the Organizations Index. At first I was unsuccessful due to my lack of knowledge of British corporations and organizations. My initial query was a simple search using the term “waterstone”, for Waterstone’s Books, which is a British chain of booksellers that I am familiar with. However, no results were returned. Under the search field on the “Simple Search” page is a box which contains links to more advanced means of searching, including separate searches of the two indexes listed above, as well as an alphabetical search which lists all of the businesses that are indexed. Each letter of the alphabet is subdivided according to the first two letters of the organization name: businesses that begin with the letters “Ab”, “Ac”
and so forth. The section for “Ab” alone had 516 entries. I also did a search under “Un”, which had 759 entries to see if I could find the University of Kent at Canterbury where I did an exchange program. There was no entry, but I did find one for “Un-Named Junk” whose name intrigued me. Clicking on it led to a page that had tabs for “Archival Information” and “Historical Information”. The scope of the entry was “1618: journal of Edmund Sayers, merchant” and could be found in the British Library: Asia, Pacific and Africa Collections (previously Oriental and India Office Library).

For the evaluation of the OAC, I did a query using the term “gold rush” as it is a key event in California History meaning that I would have a greater likelihood of a successful query. The results page contained two tabs, one of which listed all of the collections (416 collections) and the other of which listed all the actual items (913 items). Each listing contained the title of the collection as a hyperlink, beneath which was the name of the contributing institution as well as a brief description of the collection. Search terms are provided beneath each listing where available. A sidebar to the right of the listings contains links to the contributing institutions as well as the number of collections each institution has. Users can also filter by the date range of the materials within the collection, and can also view the online items that are available. Clicking on the collection name brings you to a summary of the collection including location, the collection number and whether or not it has any items available online. The “Items” tab contains direct links to digitally available materials. Some materials such as maps and photographs are direct scans, though some items, such as letters and other written records are only available in transcript form, though they may be accompanied by related scanned items.

For the NUCMC I conducted an advanced search using the term “lincoln, abraham” as “Author—Personal”, I located 751 records. I selected the first record and followed the instructions on the link “Locating Collections” that was on the NUCMC page. The instructions required that I click “More on this Record” at the bottom of the listing, which led me to an expanded record that looked like a standard library card catalog entry. At the bottom of this page I clicked on the link for “Tagged display” which took me to the MARC21-encoded entry. According to the instructions, I looked at the 040 field, and entered the letter code on the Participating Institutions Search page (OCLC, 2010a) which was a separate website I first entered the code in the field for “OCLC Symbol” and then in the field for “MARC21 Organization Code”, as I was not familiar with MARC21 and thus did not know where the code should be entered. However, the entry had multiple letter codes, each of which I entered in the “MARC21 Organization Code” field, and which provided me with numerous listings of various institutions that might have this item.

The Repositories of Primary Sources is a directory of archives, and has no search feature. The user must have an idea of where he or she would be most likely to find the item in question.

Comparative Strengths And Weaknesses

I was impressed with WorldCat’s comprehensiveness. The fact that users have access to 1.5 billion items worldwide (OCLC, 2010) makes this a key resource for researchers. The sidebar feature in the results page provides a lot of information to the
user about the various formats and features of the holdings that are available based on the query. However, I found that the user has to be persistent in drilling down. I only figured out that the result for my query on “Halldor Laxness” was a hyperlink by guessing, and clicking on it. Only after doing so was I brought to the next page which gave a description of the collection and let me search for the library that held it. I feel that the information presented on the results page could have consisted of the hyperlink only (without the list of people underneath). This would be especially useful if there are numerous results returned, meaning that the user would not have to wade through a lot of text to find what he or she is looking for. Another weakness is that there is no listing available of the participating libraries.

One feature of the National Register of Archives that I found useful was the fact that users could perform both simple and advanced searches. In addition, users could perform searches of single indexes, or alphabetical lists where available. I found the alphabetical list the most helpful when searching organizations and corporations. The NRA site also offers a comprehensive “Help” section that contains searching tips for the various indexes, a FAQ section, and a glossary. The home page of the NRA also contains research guides for various topics. One problem I found is that the only direct link between the National Archives and the NRA is on the National Archives home page. The NRA is not linked from any of the National Archives’ other pages, meaning that a user has to navigate back to the home page in order to get back to the NRA. The other option is to use the A-Z Index that is located at the bottom of every page of the National Archives site. Seeing as the NRA acts as the online catalog to the National Archives, research would be facilitated by having a link to the NRA located on every page of the National Archives site. Still, I am impressed by a site that can claim to index 1,000 years of government records as well those of families and estates stretching back hundreds of years.

I encountered much difficulty in trying to figure out how to use the NUCMC system. At first I tried using the “Find” field at the top of the page, in “NUCMC Web Pages”. I used the search terms “mark twain”, “thomas Jefferson”, “jefferson, thomas*” and “lincoln, a*”. Each search took me to a Library of Congress page that contained a box with recommended links, followed by a list of links that about the NUCMC, but no listings of manuscript collections. I looked at the links on the NUCMC page under the section “Search Manuscripts”, and even tried doing a direct search using the Library of Congress catalog. As I became more frustrated I used the “Ask a Librarian” feature to ask how to search NUCMC. As of the time of the writing of this paper I am still waiting for a response. I finally revisited the link for “Searching on OCLC WorldCat” and was successful. I feel that relying on the MARC21 system to locate manuscripts is unnecessarily cumbersome, considering that much more user-friendly systems exist, such as providing direct hyperlinks to collections as with OAC and WorldCat: UC Berkeley Libraries. I think that having a simplified system would aid researchers in that they would be able to find what they are looking for much more quickly and efficiently. The NUCMC website could also be re-designed to make the searching instructions more prominent and to remove the search field at the top of the page, which I found to be misleading.

I found the Repositories of Primary Sources to be a very interesting idea. It is helpful to have a directory of archives and special collections throughout the world.
However, the site is limited by the fact that it is only a directory, and that users cannot search the website as a whole for the location of particular items. Most researchers already have an idea of what they are looking for, and just want to know where they can find it. They want to be able to find the item and not have to hunt through pages of websites to find the location that may have it. Furthermore, a number of the links on the site are no longer active: I counted at least five, by doing a random sample. The home page for the Repositories of Primary Sources states that “All links have been tested for correctness and appropriateness” and provides a form to submit questions, new links or updates.

Summary

The sites described above used a variety of aids to enable to researchers to find information about collections. Some sites were more straightforward to use, such as the Online Archive of California, which I found to be the easiest to use, while at least one (the National Union Catalog of Manuscript Collections) relied on more technical classification systems, such as MARC21, which has the potential to hinder research by those who are not familiar with the format, or who do not wish to take the time to learn how to use it.

Conclusion

It is important that researchers be able to maintain their focus on the goal of their work and not have to worry about the design and accuracy of sites that provide online access to information about archival collections. As we move forward in the age of Web 2.0, it will become easier to combine cataloging techniques with easy-to-use interfaces that allow users to access a wide array of resources. Given the growing amount of materials and data in collections around the world, it will become increasingly important to make sure that this information is readily available to researchers and historians.
References


